BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 2: e26003
doi: 10.3897/biss.2.26003

OPEN ACCESS

Conference Abstract

# Challenges For Implementing Collections Data Quality Feedback: synthesizing the community experience

Deborah L Paul‡, Nicole Fisher§

‡ Florida State University, Tallahassee, United States of America
§ CSIRO National Research Collections of Australia, Canberra, Australia

Corresponding author: Deborah L Paul (dpaul@fsu.edu)

Received: 19 Apr 2018 | Published: 13 Jun 2018

Citation: Paul D, Fisher N (2018) Challenges For Implementing Collections Data Quality Feedback: synthesizing the community experience. Biodiversity Information Science and Standards 2: e26003. https://doi.org/10.3897/biss.2.26003

## Abstract

Much data quality (DQ) feedback is now available to data providers from aggregators of collections specimen and related data. Similarly, transcription centres and crowdsourcing platforms also provide data that must be assessed and often manipulated before uploading to a local database and subsequently published with aggregators. In order to facilitate broader DQ information use aggregators (GBIF, ALA, iDigBio, VertNet) and others, through the TDWG BDQ Interest Group, are harmonizing the DQ information provided - transforming part of the DQ feedback process. But, collections sharing data face challenges when trying to evaluate and integrate the information changes offered (by aggregators) for given records in local collection management systems and collection databases. Sharing DQ integration experiences can help reveal risks and opportunities. Discovering others have the same conundrums helps develop a community of belonging and may assist in removing duplication of effort. It is important to leverage the knowledge and experience of those who are currently validating data to improve the efficiency and effectiveness of the process. Documenting and classifying these challenges also facilitates motivation and community building by informing those who would tackle these challenges. In this case, talks from aggregators and data providers give all of us a chance to learn from their stories about implementing and integrating DQ feedback.

© Paul D, Fisher N. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Following the symposium, a special interest group (SIG at SPNHC) meeting offers everyone an opportunity to add their experiences with aggregator DQ feedback. See the SIG meeting: "Add Your Input to Challenges for Implementing Collections Data Quality Feedback: synthesizing the community experience", for details. For tractable issues, we plan to assemble and note the expected ways in which these barriers can be overcome. Where possible, we can tap into existing community resources (SPNHC, TDWG, biology.stackexchange.com, iDigBio, etc.) to help our data providers implement future data updates and track changes. At the same time, we plan to analyze the intractable issues - documenting why they remain challenging - and what, if any potential solutions are available or likely to be available in the future. This information provides future projects like DiSSCo (Distributed System of Scientific Collections) and BCoN (Biodiversity Collections Network) and others worldwide the information required to plan more effectively for cyber/human infrastructure. Synthesizing this input helps visionaries better understand, anticipate and support DQ management and data mobilization efforts going forward by informing design of future proposals and global projects structured with these outcomes in mind. At the end of the workshop, we intend to publish our findings, and merge them with the results of a global survey on the same topic.

## Keywords

Data quality, Data transformation, Symposium, Workshop, Data integration, Capacity building, Community building, Collections data management, Feedback assessment

## Presenting author

Deborah Paul and Nicole Fisher as Symposium and Workshop Organizers

## Acknowledgements

Support for this symposium comes from iDigBio via the National Science Foundation's Advancing Digitization of Biodiversity Collections Program (Cooperative Agreement EF-1115210), and from CSIRO Digital Collections and Informatics National Research Collections Australia (NRCA). We also note the efforts of the TDWG BDQ Interest Group and especially Arthur Chapman, in assuring that our conference activities on BDQ at SPNHC are coordinated and in concert with those at TDWG.

## Funding program

Advancing Digitization of Biodiversity Collections Program (National Science Foundation's Cooperative Agreement EF-1115210).